Getting to Ground Truth: Integrating Social and Spatial Statistics to Support our Communities' Most Vulnerable Children and Families

Christopher Church,JD[1], Melissa Strompolis, PhD[2] and Megan Branham, LMSW[2]
Children's Law Center[1]
Children's Trust of South Carolina[2]

Introduction

The phrase *ground truth* is military slang to describe the attempt to reconcile information contained in intelligence reports with the front line reality in military actions. The importance of getting to *ground truth* cannot be overstated: front line personnel require a clear understanding of the expectations, intended outcomes, and challenges they are facing in order to achieve the leadership's goals and objectives. The critical importance of getting to *ground truth* exemplifies the work our state agencies and local civic organizations undertake to strengthen children and families. State agencies and local civic organizations need a clear understanding of the problem, intended outcomes and barriers to effectively help all children reach their full potential. Understanding a problem from a national or state level vantage point disguises variations within a problem area at a local level.[1] This paper will explore a technique in getting to *ground truth* with the integration of spatial statistics and displays of minimally aggregated census data as a pseudocolor overlay on Google Maps™. The first section describes a current issue – poverty in South Carolina – and provides national, state and county-level data. The second section discusses the technical aspects of the data visualization tool. The third section provides an overview of how a member of the public can access and view the maps. Finally, the paper concludes with a discussion of how the maps can be used to support data-driven conversations and decision-making.

KIDS COUNT® and Poverty

KIDS COUNT®, an annual project of the Annie E. Casey Foundation, tracks and measures child well-being at the national and state level. Each year, KIDS COUNT releases a set of 16 indicators of child well-being to raise awareness and ignite state and local policy and practice changes to improve the lives of children.[2] (In South Carolina, KIDS COUNT is housed within Children's Trust of South Carolina.) This year, South Carolina again ranked near the bottom nationally (45th) on indicators of child well-being. Although South Carolina made improvements in some areas, indicators related to poverty were not only dismal, but also worsened from last year. Nearly 27% of children in South Carolina live in poverty (288,000 children) and about 14% of children (148,000) live in high poverty areas. The data in Table 1 show the national, state and county-level child poverty rates relevant to South Carolina.

---

[1]. Casey Family Programs. (2014). A declaration of hope. Retrieved from http://www.casey.org/hope/

[2]. http://www.aecf.org/work/kids-count/

Poverty is an important indicator of child well-being as it is highly correlated with a number of important outcomes, including physical and mental health, educational achievement and economic stability.[1,3] Clearly, these outcomes represent a desired state for all adults; and therefore, poverty is an important indicator on whether or not a child will reach his or her full potential. For organizations such as Children's Trust of South Carolina, which works to prevent child abuse, neglect and injury, ensuring that all children in South Carolina have the opportunity to thrive is relevant. Effective data on poverty can inform important partnerships and work.

TABLE 1
Percentage of Children in Poverty

Percentage of Children in Poverty in the United States and South Carolina State and Counties					
United States	23.0%	Colleton	37.5%	Lee	40.1%
South Carolina	26.8%	Darlington	34.7%	Lexington	21.4%
Abbeville	28.9%	Dillon	41.0%	Marion	44.3%
Aiken	27.4%	Dorchester	16.9%	Marlboro	35.7%
Allendale	50.6%	Edgefield	27.6%	McCormick	36.3%
Anderson	24.0%	Fairfield	34.9%	Newberry	27.0%
Bamberg	40.3%	Florence	29.1%	Oconee	25.3%
Barnwell	40.0%	Georgetown	35.0%	Orangeburg	38.0%
Beaufort	22.1%	Greenville	23.2%	Pickens	20.8%
Berkeley	20.9%	Greenwood	30.4%	Richland	23.0%
Calhoun	30.8%	Hampton	38.4%	Saluda	28.7%
Charleston	28.1%	Horry	32.5%	Spartanburg	27.2%
Cherokee	33.0%	Jasper	39.4%	Sumter	31.2%
Chester	36.5%	Kershaw	25.0%	Union	32.9%
Chesterfield	35.6%	Lancaster	29.2%	Williamsburg	41.7%
Clarendon	44.8%	Laurens	35.0%	York	17.2%

Courtesy of KIDS COUNT®

These data shed light on the rueful reality that exists in many of South Carolina's communities. However, they also beg the question of how state agencies and local civic organizations can use data that are presented in the context of geographical boundaries, like the county-level child poverty data in Table 1. For example, Richland County has a total area of 772 square miles.[4] Are the 23% of children living in poverty spread evenly amongst those 772 square miles? Or are they concentrated in a handful of neighborhoods? All too often, data that are meant to ignite action to address an alarming

3 The Annie E. Casey Foundation. (2014). *Creating opportunities for families: a two-generation approach.* Retrieved from http://www.aecf.org/resources/creating-opportunity-for-families/

4 Richland County South Carolina. (n.d.). In *Wikipedia*. Retrieved December 10, 2014, from http://en.wikipedia.org/wiki/Richland_County,_South_Carolina

problem do not provide enough information to act where help is most needed. The data visualization tool discussed in this paper was developed to explore and answer questions like these. With minimally aggregated spatial data visualized in a meaningful way, our state agencies and local civic organizations can identify social problems, such as child poverty, at the precision of a street address or apartment complex. This precision enables organizations to meaningfully target resources and interventions where the data tell us they are most needed.

Technical Discussion

Organizations working with children and families have access to a wealth of demographic, service, outcome and other data. These data may come from the organization's own collection efforts, community partners or public sources such as KIDS COUNT. The data are used for a variety of reasons, and displayed in a variety of ways. A common approach is to display data geographically. For example, the choropleth map in Figure 1 displays the percent of children living in poverty in South Carolina, by county.[5]

FIGURE 1
Choropleth Map of Child Poverty (%) in South Carolina by County, 2012
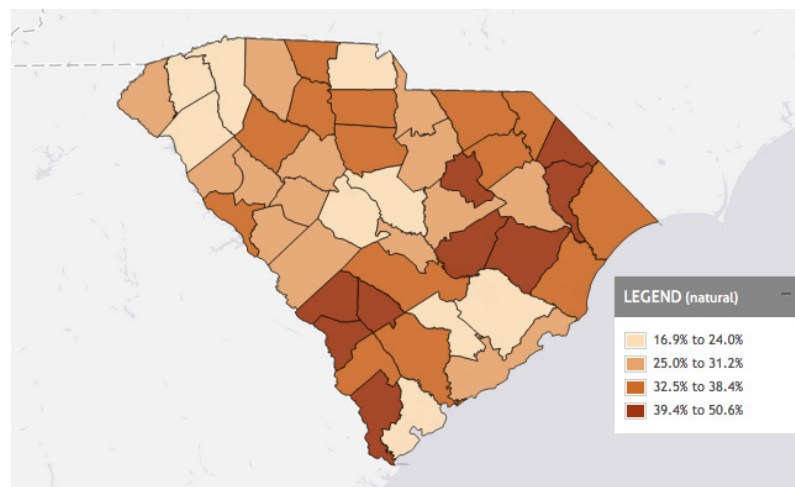
Courtesy of KIDS COUNT®

A choropleth map is a geographical data visualization tool that displays different colored polygons to represent differences across geographies. For example, in Figure 1, the polygons represent the different counties in South Carolina, and the color shading scheme represents a pre-defined relationship between one of the colors and the percent of children living in poverty in a given county.

While choropleth maps have utility, they also introduce significant statistical bias when the polygon does not bear any meaningful relationship to the problem examined. This bias is commonly referred to as the modifiable areal unit problem, which occurs when point-level spatial data are aggregated into larger polygons such as zip codes or counties.[6,7] Aggregating geospatial data into larger polygons makes the data even more difficult to interpret.

---

[5] United States Census Bureau - Small Area Income & Poverty Estimate [SAIPE]. (2012). *Under age 18 in poverty.* Retrieved from http://tinyurl.com/SCchildpoverty
[6] Openshaw, S. (1983). *The modifiable areal unit problem* (Vol. 38). Norwich: Geo Books.
[7] Wong, D. W. (2009). *The modifiable areal unit problem* (MAUP). In A.S. Fotheringham & A. Stewart (Eds.), The SAGE handbook of spatial analysis. (pp. 105-124). London: Sage Publications.

For example, compare the two maps in Figure 2, displaying child poverty in an area of Atlanta, Georgia near the intersection of Interstate 20 and Interstate 285. In the map on the left, the child poverty rate is aggregated into zip code polygons, which can contain over tens of thousands of people.[8] In the map on the right, referred to as a heat map, the child poverty rate is aggregated into census blockgroup polygons, which most often contain between 600 and 3,000 people.[9] In comparing the two maps, it is clear how aggregating data to larger geographies can lead to data misinterpretations, such as hiding clustering and using deceiving, non-neutral polygons, as explained by the modifiable areal unit problem. In other words, as is the case in Figure 2, a single blockgroup with a high poverty rate can influence a zip code level (or other larger geographical area) choropleth map, which can lead to interpretation errors of the underlying spatial data.

FIGURE 2
Comparison of Choropleth and Heat Map at Interstate 20 & 285 Intersection in Atlanta, GA
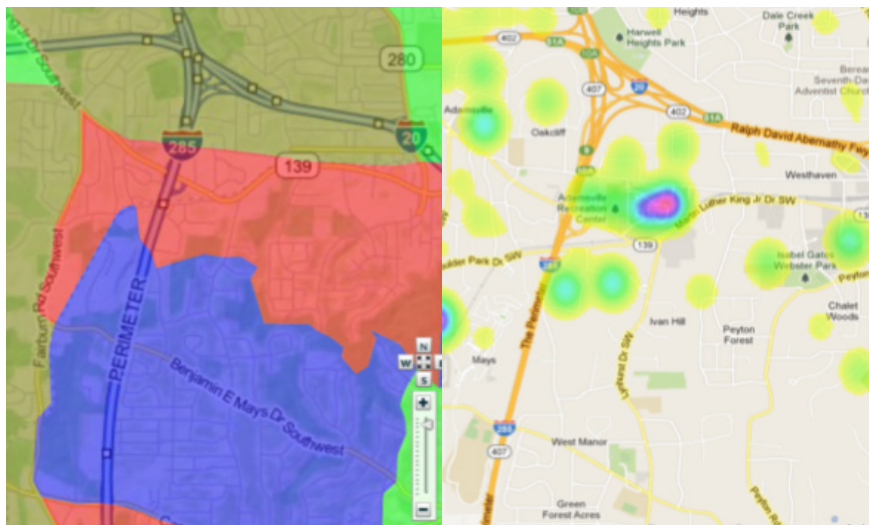
Courtesty of Fostering Court Improvement

Minimally aggregated spatial data identify problems with more precision, and enable a deeper understanding of where certain phenomena, such as child poverty, occur in a community. The map in Figure 2 on the right is the data visualization tool discussed in this paper, and the data source is publicly available United States Census Bureau data.

Heat Maps

Heat maps are created by grouping the census data at the block or blockgroup centroid using binned kernel density estimation. The data are then clipped onto a 1024x1024 grid covering the county for which the data are contained, and smoothed with a Gaussian kernel to create a transparent pseudocolor overlay aligned to Google Maps™. This process allows users to view minimally aggregated spatial data for the purpose of examining the relationship between certain social and demographic clusters of data at precise geographical locations within a county.

---

[8] Stiles, M. (2013, July 01). The Zip Code Turns 50 Today; Here are 9 That Stand Out. *National Public Radio*.
Retrieved from http://www.npr.org/blogs/thetwo-way/2013/07/01/197623129/the-zip-code-turns-50-today-here-are-9-that-stand-out
[9] United States Census Bureau. (2010). Geographic terms and concepts – block groups.
Retrieved from https://www.census.gov/geo/reference/gtc/gtc_bg.html

The maps are created by geocoding two forms of census data: the Decennial Census of Population and Housing from 2010 and the American Community Survey (ACS) for the aggregated years of 2008-2012. The heat maps for South Carolina contain the census data aggregated at the block count level (see Table 2).

The ACS collects demographic, economic, programmatic and social data annually.[10] These data are released periodically, often in multi-year datasets. The heat maps for South Carolina contain ACS statistics aggregated at the blockgroup count level.

TABLE 2
Decennial Census & ACS Heat Maps Available in South Carolina

Data Source	Indicator
Decennial Census	Child Population Aged Zero to Four
	Child Population Aged Ten to Seventeen
	Child Population
	Child Population Living in Non-White Householder
	Child Population Living in Hispanic Householder
	Child Population Living in Single Parent Householder
	Child Population Living in Grandparent Householder
	Persons Living in Renter-Occupied Household
	Persons Living in Renter-Occupied Five or More Person Household
American Communities Survey	Families with Children Below Poverty Line
	Household Aggregate Cash Public Assistance
	Persons Aged 25 or Older with No High School Education
	Families with Children and No Employment Parent

Accessing the Maps

Maps for every South Carolina county are available at: **www.fosteringcourtimprovement.org/maps/SC.**

To access the maps, visit the website above. A page, as seen in Figure 3, contains each county name as a hyperlink.

To access the county map, simply click on the county name. For example, by clicking on "Aiken", a webpage containing side-by-side maps of Aiken County will open, as seen in Figure 4.

The webpage contains certain defaults for the maps. To view any of the heat maps generated from census or ACS data, simply click the drop-down menu above the map and select the desired data element. In the lower left-hand corner of the webpage, two selections may be made. The first is the *opacity*, which can be changed to increase or decrease the transparency of the pseudocolor overlay.
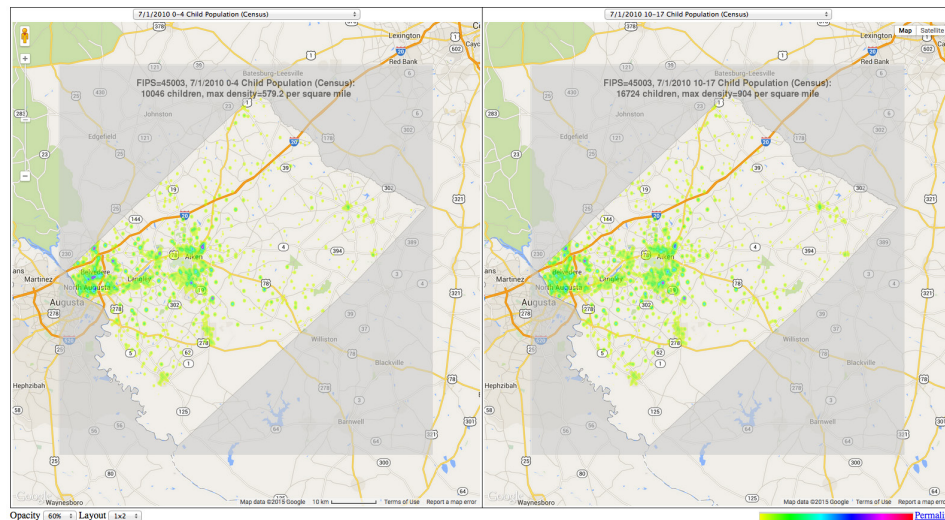
---

[10] http://census.gov/acs/

FIGURE 3

Index of Maps for South Carolina Counties

Courtesy of Fostering Court Improvement

FIGURE 4

Aiken County Heat Map

Courtesy of Fostering Court Improvement

The second is the *layout*, which can be adjusted to change the number of maps (from 1 to 9) displayed on the screen. The layout is in the *Row x Column* format. Since the data are displayed as a transparent intensity overlay aligned with Google Maps™, features such as zooming and Street-View may be used to explore the maps.

Finally, clicking on the "Permalink" hyperlink in the bottom right-hand corner of the page allows the user to share the maps. The permalink will share the map while preserving the specific settings related to the number of maps, data source, opacity and level of zoom.
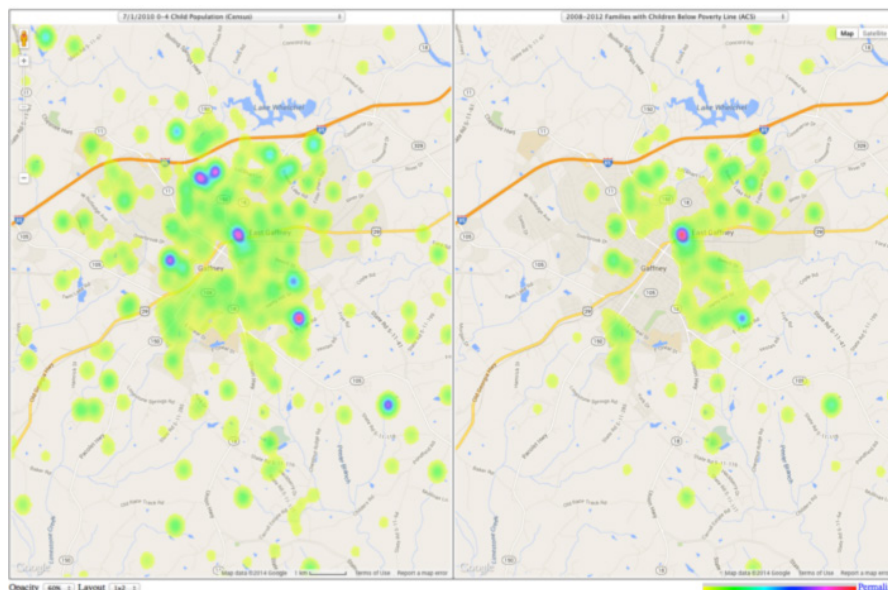
Using the Maps

The maps are useful in a variety of practice, planning and policy and advocacy contexts. In many governmental and civic organizations, a cornerstone of quality improvement work requires the use of data to enhance understanding within the nature of the work. The application of localized data would not only improve decision making at all levels, but also capitalize on the strengths of local communities to support their most vulnerable members.[11]

These novel maps are an engaging way of displaying administrative data. Additionally, local data increases the relevancy and interest to the data consumer. Not only would a staff member from a Columbia-based nonprofit find Richland County child poverty rates more relevant compared to South Carolina child poverty rates, that staff member would likewise find zip-code level data more meaningful than county-level data. The more local the data are, the more meaningful they will be to the user. The data displayed in these maps are the most localized data publicly available, often displaying concentrations of social and demographic data at the precision of a large apartment complex or the neighborhood level. These maps are relevant and useful for students, professionals and the public-at-large to explore and enhance their understanding of their communities.

Similarly, the heat maps are useful for planning purposes. For example, consider the map in Figure 5 of Gaffney, South Carolina. The map on the left shows the population of children ages zero to four and the map on the right shows the number of children living in poverty. The Gaffney community might decide that given the large number of young children living in poverty, that a public Pre-K school near the intersection of Highway 29 and state road 150 would be an effective way to serve that population.

FIGURE 5
*Heat Map of Young Children and Children Living in Poverty in Gaffney, SC*

Courtesy of Fostering Court Improvement

---

[11] Casey Family Programs. (2014). A declaration of hope. Retrieved from http://www.casey.org/hope/

Along the same lines, the maps are an invaluable tool for enhancing the decision making process as it relates to allocation of public and private resources. For example, Figure 6 displays the concentration of victims of child abuse and neglect for Savannah, Georgia during the 2010 calendar year.

FIGURE 6
Heat Map of Substantiated Child Abuse and Neglect Cases in Savannah, GA
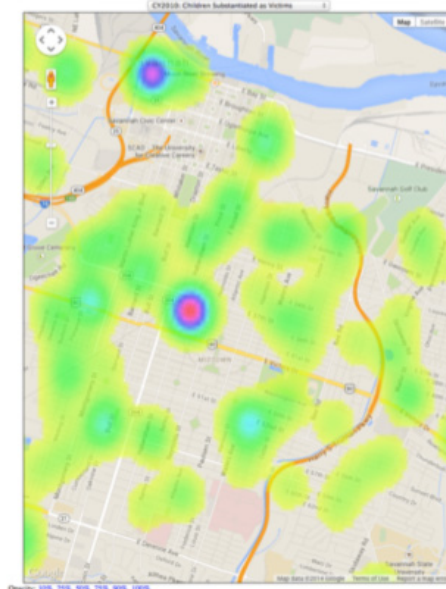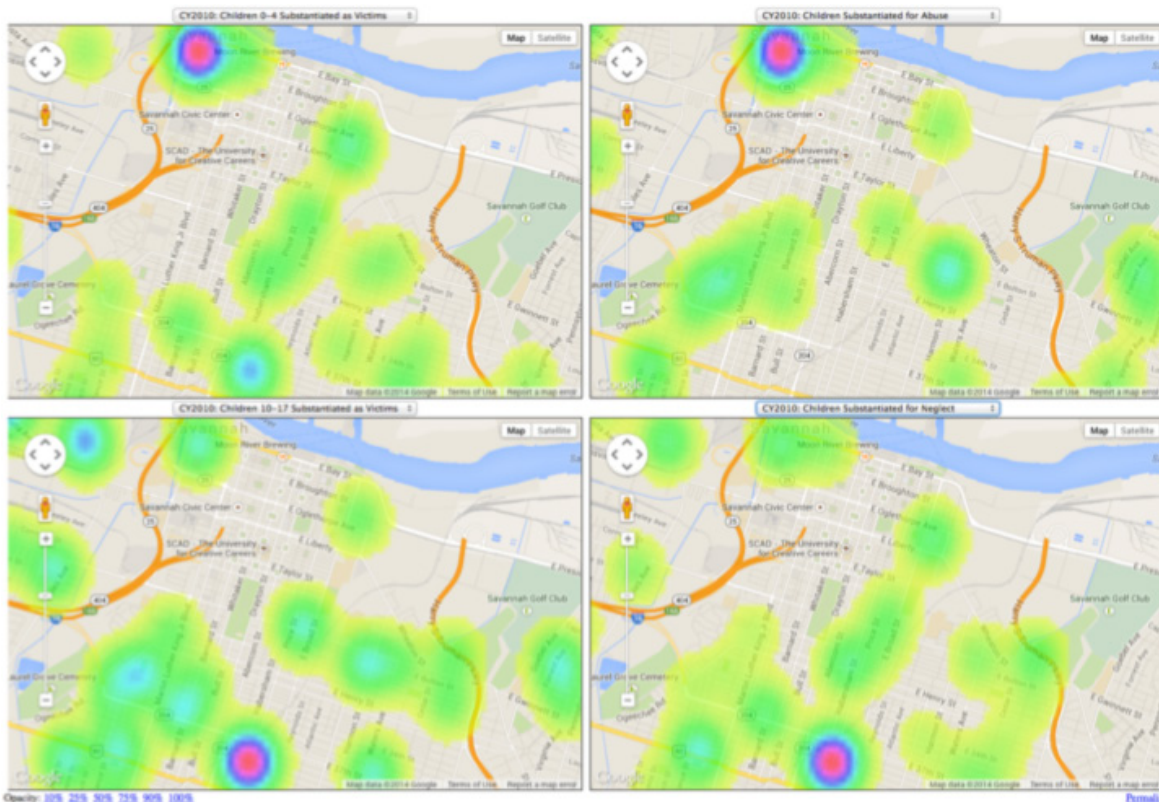

Courtesy of Fostering Court Improvement

There are two spots on the map that reveal intense concentrations of child abuse and neglect victims. However, an examination of the data stratified by age and type of abuse (primary allegation) reveals a stark contrast between the two neighborhoods.

FIGURE 7

Heat Map of Substantiated Child Abuse and Neglect Cases by Age and Allegation in Savannah, GA

Courtesy of Fostering Court Improvement

As shown in Figure 7, the neighborhood on the north end of town has a concentration of young victims (age zero to four; upper left quadrant) where the primary allegation was child abuse (upper right quadrant). The neighborhood on the south end of town, by contrast, shows a high concentration of older children (age ten to seventeen; lower left quadrant) where the primary allegation was child neglect (lower right quadrant). Certainly, these two neighborhoods will require very different public and private resources and services to ensure children in those neighborhoods have the opportunity to reach their full potential.

The example discussed in Figure 7 also highlights another important use of the maps: policy and advocacy. Although the two neighborhoods in Savannah, Georgia, face different challenges, they may require similar advocacy and policy strategies. Complex social challenges, including child abuse and neglect, require policy solutions that bolster systems that support families and strengthen communities. These challenges, and many times solutions, are not unique across zip codes, but sharing of these data can better align advocacy efforts. All communities play an important role in identifying, advocating for and supporting programs for families at high risk of abuse and neglect.[12]

Heat maps allow community leaders and policy makers the opportunity to clearly identify and thus respond effectively to child abuse and neglect. Additionally, the heat maps provide a powerful visual

---

[12] State Policy Advocacy and Reform Center. (2014). *Child welfare policy primer: a guide for advocates, policymakers, and others interested in child welfare policy reform*. Retrieved from http://childwelfaresparc.org/wp-content/uploads/2014/06/Child-Welfare-Policy-Primer-Final2.pdf

to track changes over time. For example, it may be appropriate for city and county elected officials to discuss affordable, safe housing and public transportation to meet family needs. Additionally, discussions with state elected officials related to equitable family tax and income policies and state appropriations for evidence-based parent education programs may be relevant. Ultimately, heat maps should spur aligned stakeholder conversations with city, county and state elected officials to achieve systems-level solutions to meet local challenges.

Cautions

Heat maps are a powerful tool for identifying needs and gaps at very localized levels. That type of identification, however, must be accompanied with use that is culturally competent and appropriate. For example, the Brookings Institute and University of Wisconsin-Milwaukee Employment and Training Institute issued a discussion paper on the purchasing power of central city neighborhoods.[13] In the discussion paper, the authors note that use of economic data by marketing firms resulted in racial and class-based stereotype statements. While the inappropriate use of economic data by marketing firms may not be the norm, any activities resulting from the heat maps must be thoughtfully considered and appropriate for the data used. Furthermore, the use of heat maps should be part of a larger, strengths-based approach for working with local communities.

"Strengths-based approach is a manner of working with individuals, families, and organizations grounded in the principle that those individuals: have existing competencies; have resources; are capable of learning new skills and problem-solving; can use existing competencies to identify and address their own concerns; and can be involved in the process of healing and self-health. It is based on the idea that people and environments interact and change each other in the process. Each has the ability to build the other's capacity"(p.150).[14]

Conclusion

Although there are risks to using heat maps to identify areas of need, the benefits greatly outweigh the risks. As stated by Jolie Bain Pillsbury in *Theory of Aligned Contributions: An Emerging Theory of Change Primer*[15], "The visceral reactions of surprise, compassion, dismay, dissatisfaction, and sometimes outrage upon realizing and seeing the state of children, families and communities create urgency. Population level data including graphs in the form of…geomaps make the current condition visible and accessible. Disaggregated data further illuminates the current condition by providing information about race/ethnicity, socioeconomic status, gender, geographical districts, special needs populations and other descriptive data relevant to the problem" (p. 6). The heat maps provide not only a conversation starter for neighborhoods and communities, but also locations for targeted programs, services and policy and advocacy efforts.

Acknowledgement

[13] Pawasarat, J., & Quinn, L. M. (2001). *Exposing urban legends: The real purchasing power of central city neighborhoods.* Washington, DC: Brookings Institution, Center on Urban and Metropolitan Policy.

[14] Ponnuswami, I., Francis, A., & Udhayakumar, P. (2012). Strengths-based approach to social work practice with older persons. In P. Venkat, L. Chenoweth, A. Francis, & S. Bakaj, (Eds.), Papers in strength based practice. (pp. 150-161). New Delhi: Allied Publishers.

[15] Pillsbury, J.B. (n.d.). *Theory of aligned contributions: an emerging theory of change primer.* Sherbrooke Consulting Inc. Retrieved from http://www.sherbrooke-consulting.com/products/TOAC.pdf